The accredited security level of this system is:

TOP SECRET // SI-GAMMA / TALENT KEYHOLE // ORCON / PROPIN / RELIDO/ REL TO USA, FVEY *

TOP SECRET//SI//REL TO USA, FVEY

[edit] (U) SSO Collection Optimization Overview

(S//SI//REL USA, FVEY) This wiki article collects and documents the activities within SSO to optimize collection from SSO sites. It is a direct result of the activities and findings from the Large Access Exploitation (LAE) Working Group. Optimization activities can be categorized as: impacting primarily content collection, i.e., feeding repositories like PINWALE and PRESSUREWAVE; metadata, perhaps better termed "structured data," collection, i.e., primarily feeding repositories like MARINA or MAINWAY; or both.

(TS//SI//REL USA, FVEY) An examination into the content collected by SSO sites in the fall of 2011 revealed that a significant portion of collection was repetitive, better placed into metadata repositories, or of little foreign intelligence value. Rapidly changing internet protocols, imprecise targeting methods, and constantly shifting target technology use means that selectors or traffic seen by tasking today may not be the same tomorrow. In addition, several emerging technologies in use by targets or contacts of targets have protocols which can cause gross over-collection and selector detasking, e.g., Yahoo! Webmessenger, which cannot be prevented under UTT tasking, as the protocol contains the precise selector tasked.

(TS//SI//REL USA, FVEY) The SSO Optimization team's job is to identify these types of data, and ensure appropriate corrective action is taken, throttling the data from corporate content or metadata repositories, as appropriate.

(TS//SI//REL USA, FVEY) Implementation details for specific SSO sites can be found on the SSO Collection Optimization NOFORN wiki page.

~~~~~~
~~~~~~

[edit] (U) Address Books

(U) Reason for Optimizing

(TS//SI//REL USA, FVEY) Ownerless address books can account for 22% 2012), are highly repetitive, and can be both numerous and large. Additionally, the data is highly structured, so they are better parsed and analyzed in structured repositories like MARINA.

(U) Criteria (SCISSORS)

(TS//SI//REL USA, FVEY) To alleviate the large number of ownerless address books currently being collected, address books will be blocked at SCISSORS. SCISSORS will process each flow and label sessions with the category of 8223 where they meet the following criteria:
• (Yahoo|Gmail|Hotmail)-viewAddressBook (which represents ~90% of all addressbook collection)
• APMSGTYPE IS PRESENT
• NO APACTIVEUSER
• SIGAD = Given SIGADS (currently, four)
(TS//SI//REL USA, FVEY) The sessions that do not meet the above criteria will send metadata to FALLOUT, but the sessions will be dropped before reaching PINWALE. Metrics for category 8223 will not be available in YELLOWSTONE, and these volumes will not count against PINWALE site caps set by the CSRC.

(U) Criteria (XKEYSCORE)

(TS//SI//REL USA, FVEY) Because of how inefficient the SCISSORS implementation of ownerless address book throttles is (essentially requiring that SCISSORS process all of the selected content from the affected dataflows twice), work is underway to implement the same throttle in XKEYSCORE at site. This requires a software modification to allow certain types of ownerless buddy list metadata to be created at site (vs at SCISSORS, under the current architecture). Then a deployed XKEYSCORE fingerprint will block the content while allowing the metadata to be memorialized.

XKS Versions:
XKS Version 1.5.8 v89 TU Version 1.5.7 v151

Digester Labeling:
<HASBUDDY>true</HASBUDDY>
<ANONADDRBOOK>true</ANONADDRBOOK>

The following fingerprint labels and defeats the traffic.
defeat/atxks/ownerless_addressbook

(U) Deployment Date

(TS//SI//REL USA, FVEY) The SCISSORS ownerless address book throttle was implemented on 2/29/2012 for one site, and for others in March and April of 2012.

(TS//SI//REL USA, FVEY) The XKEYSCORE ownerless address book throttle is under development.

~~~~~~

[edit] (U) Yahoo Webmessenger

(U) Reason for Optimizing

(S//SI//REL USA, FVEY) Yahoo's web-based Messenger client sends frequent requests, and receives frequent responses, for inbox and buddy status which are highly repetitive and contain little or no useful FI information in the actual message content beyond the simple fact that the user was online. During the first two weeks of December, at least eight selectors were detasked by CSRC due to excessive collection (exceeding session limits).

(U) Criteria
~~~~~~

(S//SI//REL USA, FVEY) Yahoo Webmessenger is identified as either being a request for status message from the client (uncommon), or a status message (common):

• request

1. HTTP GET /v1/pushchannel/
2. 'x-yahoo-msgr-user-agent: YahooMessenger'
3. http_host('rest-notify.msg.yahoo.com')
4. http_url('&msgrAppId=' and '&sid=')

• status

1. 'Content-Type: application/json;charset=utf-8'
2. Message contains: '"@pendingMsg" : 0, "@syncStatus" : 0, "responses" : [{'

(U) Deployment Date

(S//SI//REL USA, FVEY) /atrouter/ on 11 Jan 2012. /atxks/ on 17 April 2012.

~~~~~~
~~~~~~